MA 3046 - Matrix Analysis

Problem Set 2 - Section I - Fundamenals

1. Show that, for any nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{C}^m$, the product $\mathbf{u}\,\mathbf{v}^H$ is a rank one matrix in $\mathbb{C}^{m\times m}$. Also show that, if $\mathbf{v}^H\,\mathbf{u} \neq -1$, then

$$\mathbf{I} - \frac{\mathbf{u}\,\mathbf{v}^H}{\mathbf{v}^H\mathbf{u}+1} = \left(\mathbf{I} + \mathbf{u}\,\mathbf{v}^H\right)^{-1}$$

(Note that $\left(\mathbf{I} + \mathbf{u}\,\mathbf{v}^H\right)$ is commonly called a *rank one perturbation of the identity.*)

2. Consider the following matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix} \quad , \quad \mathbf{B} = \begin{bmatrix} 1 & 2 & -1 \\ -2 & 1 & 1 \\ 1 & -2 & 1 \end{bmatrix} \quad , \quad \mathbf{C} = \begin{bmatrix} \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{2} \\ \frac{3}{6} & \frac{1}{2} \\ \frac{5}{6} & -\frac{1}{2} \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \quad , \quad \mathbf{E} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{i}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{i}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{i}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{i}{2} \end{bmatrix}$$

where $i = \sqrt{-1}$. Which, if any, of these are unitary? Which, if any, are orthogonal? Of those that are neither, which are *easily* converted to unitary ones?

3. Consider the matrix:

$$\mathbf{C} = \begin{bmatrix} \frac{1}{6} & \frac{5}{6} \\ \frac{3}{6} & -\frac{1}{6} \\ \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & \frac{3}{6} \end{bmatrix}$$

Show that $\mathbf{C}^H\mathbf{C} = \mathbf{I}$, but $\mathbf{C}\mathbf{C}^H \neq \mathbf{I}$ and so $\mathbf{C}^H \neq \mathbf{C}^{-1}$. Explain why this is not a contradiction.

4. Consider the matrix-vector equation

$$\mathbf{B}\,[\,\mathbf{x}\,]_{\mathbf{B}} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} [\,\mathbf{x}\,]_{\mathbf{B}} \;=\; \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix}$$

a. Solve this system for $[\,\mathbf{x}\,]_{\mathbf{B}}$ by Gaussian elimination.

b. Solve this system for $[\,\mathbf{x}\,]_{\mathbf{B}}$, without using elimination, but using the facts that $\mathbf{B}$ is unitary and $[\,\mathbf{x}\,]_{\mathbf{B}}$ represents the coordinates of $[\;1\;\;-7\;\;\;2\;]^H$ in terms of the columns of $\mathbf{B}$.

5. Show that if $\mathbf{Q}^{(1)}$ and $\mathbf{Q}^{(2)}$ are **any** two unitary matrices of the same size, then their product, i.e. $\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}$ is also Unitary.

6. Consider the three most common measures for the norm of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ (or the equivalent row or column vector forms):

$$\begin{aligned} \|\,\mathbf{x}\,\|_1 \; &= |x_1| + |x_2| + \cdots + |x_n| \\ \|\,\mathbf{x}\,\|_2 \; &= \sqrt{x_1^2 + x_2^2 + \cdots x_n^2} \\ \|\,\mathbf{x}\,\|_\infty &= \; \max_i |x_i| \end{aligned}$$

Compute each of these norms for each of the following vectors.

a. $\mathbf{x} = (-1, 2, -2)$

b. $\mathbf{x} = (-4)$

c. $\mathbf{x} = (10, -3, 12)$

d. $\mathbf{x} = (102, -17, -1)$

e. $\mathbf{x} = (.1, -.2, -.4)$

f. $\mathbf{x} = (-12, -2, 4, 6, 5)$

7. Any vector norm *induces* a corresponding matrix norm according to the relationship:

$$\|\,\mathbf{A}\,\| = \max_{\mathbf{x}} \frac{\|\,\mathbf{A}\,\mathbf{x}\,\|}{\|\,\mathbf{x}\,\|}$$

Show that, for any matrix norm and any scalar α, $\quad \|\,\alpha\,\mathbf{A}\,\| = |\alpha|\,\|\,\mathbf{A}\,\|$

8. Consider the three matrices:

$$\text{a.}\quad \begin{bmatrix} 2 & -5 & 1 \\ 1 & 3 & 2 \end{bmatrix} \qquad \text{b.}\quad \begin{bmatrix} 10 & 7 & -2 \\ 6 & 4 & -1 \\ -2 & 1 & 1 \end{bmatrix} \qquad \text{c.}\quad \begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 1 & 2 \end{bmatrix}$$

For each of these matrices, and for each of the norms described in problem 6:

(i.) Pick five different "input" vectors ($\mathbf{x}$). (Make sure at least some of them have some *negative* components!)

(ii.) For each of these inputs, compute the ratio $\dfrac{\|\mathbf{A}\,\mathbf{x}\|}{\|\mathbf{x}\|}$

(iii.) Based on your answers to part b., determine a lower bound for $\|\mathbf{A}\|$

(iv.) Compare your lower bound to the actual corresponding value of $\|\mathbf{A}\|$ as determined by MATLAB's relevant **norm( )** command.

9. The infinity norm of a vector:

$$\| x \|_\infty = \max_i |x_i|$$

i.e., the component with the greatest magnitude, induces a corresponding matrix norm:

$$\|\mathbf{A}\|_\infty = \max_{\mathbf{x}} \frac{\|\mathbf{A}\,\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \max_{\|\mathbf{x}\|_\infty = 1} \|\mathbf{A}\,\mathbf{x}\|_\infty$$

Show that $\quad \|\mathbf{A}\|_\infty = \max_i \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}$, or, in other words, the matrix infinity norm is just the largest of the sums of the magnitudes of the coefficients on each row. For this reason, the matrix infinity norm is commonly called the *row-sum norm.*

10. Consider the following vectors and matrices:

$$\text{(a.)}\quad \mathbf{x} = \begin{bmatrix} -6 \\ 4 \\ -4 \end{bmatrix} \qquad \text{(b.)}\quad \mathbf{y} = \begin{bmatrix} -9 \\ 2 \\ -9 \\ -2 \end{bmatrix} \qquad \text{(c.)}\quad \mathbf{z} = \begin{bmatrix} -1 \\ -4 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

$$\text{(d.)}\quad \mathbf{A} = \begin{bmatrix} 8 & 8 & -4 & 2 \\ 5 & -6 & -1 & -2 \\ -7 & -4 & -9 & 0 \\ -10 & 3 & 10 & -3 \end{bmatrix} \qquad \text{(e.)}\quad \mathbf{B} = \begin{bmatrix} -6 & -10 & -2 & 7 & 0 \\ -6 & 5 & 7 & -10 & 4 \\ 2 & -1 & 1 & 4 & -1 \\ -5 & 9 & -6 & -2 & -4 \\ -6 & -1 & 3 & 7 & -6 \end{bmatrix}$$

Find the infinity norm of each.

11. Show directly that the one norm, defined by

$$\| \mathbf{x} \|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|$$

satisfies the following general properties:

(a.) $\| \mathbf{x} \|_1 > 0 \ , \quad \mathbf{x} \neq 0$

(b.) $\| \mathbf{x} + \mathbf{y} \|_1 \leq \| \mathbf{x} \|_1 + \| \mathbf{y} \|_1$, where $\mathbf{x}$ and $\mathbf{y}$ are any vectors. (This so-called *triangle inequality* essentially ensures the shortest "distance" between points must be the line connecting them.)

(c.) $\| \alpha \mathbf{x} \|_1 = |\alpha| \| \mathbf{x} \|_1$, where $\mathbf{x}$ is any vector and α is any scalar. (This formula ensures that multiplying any vector by a scalar factor simply changes its "length" by that factor, and also that $\| \mathbf{0} \|_1 = 0$.)

12. Using MATLAB, graph

$$\{ \mathbf{A}\,\mathbf{x} \mid x_1^2 + x_2^2 = 1 \}$$

for each of the following matrices. Based on your figures, estimate the singular values of each matrix, and then compare your estimates to the results of MATLAB's **svd( )** command.

a. $\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$ b. $\begin{bmatrix} -4 & -1 \\ 2 & 2 \end{bmatrix}$ c. $\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$

13. Find the singular value decomposition of each of the following matrices, and compare your results to the results of MATLAB's **svd( )** command.

a. $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}$ c. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

14. Using the singular value decomposition, show that, for any matrix $\mathbf{A}$,

$$\text{Null}(\mathbf{A}^{\mathbf{H}}\mathbf{A}) = \text{Null}(\mathbf{A})$$

15. The SVD is commonly developed from an eigenvalue and eigenvector approach instead of from the viewpoint of Euclidean norms. Specifically, we can show that, for any matrix $\mathbf{A}$, the singular values satisfy $\sigma_i = \sqrt{|\lambda_i|}$, where the λ_i are the eigenvalues of $\mathbf{A}^H\mathbf{A}$, and the right singular vectors of $\mathbf{A}$ (i.e., the $\mathbf{v}^{(i)}$) are the eigenvectors of $\mathbf{A}^H\mathbf{A}$. This leads to constructing the (reduced) SVD via the following procedure:

(i.) Form the product $\mathbf{A}^H\,\mathbf{A}$.

(ii.) Find the eigenvalues (λ_i) and eigenvectors $(\mathbf{v}^{(i)})$ of that product.

(iii.) For each $\lambda_i \neq 0$, define $\sigma_i = \sqrt{\lambda_i}$, where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > 0$.

(iv.) For each $\sigma_i \neq 0$, define $\mathbf{u}^{(i)} = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}^{(i)}$, where $\mathbf{v}^{(i)}$ is the eigenvector associated with λ_i.

(Unfortunately, as we shall see later, this procedure is not particularly well suited for “large” matrices.) Apply this procedure to find the SVD of the matrix

$$\mathbf{A} = \begin{bmatrix} 26 & 18 \\ 1 & 18 \\ 14 & 27 \end{bmatrix}$$

16. Repeat the calculation of the (reduced) SVD using the eigenvalue and eigenvector approach from problem 15 for the matrix:

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 2 & 4 \\ 2 & 4 \\ 2 & 4 \end{bmatrix}$$

17. Repeat the calculation of the SVD using the eigenvalue and eigenvector approach from problem 15 for the matrix:

$$\mathbf{A} = \begin{bmatrix} 16 & -2 \\ 13 & 14 \end{bmatrix}$$

18. Repeat the calculation of the (reduced) SVD using the eigenvalue and eigenvector approach from problem 15 for the matrix:

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$$

19. The equation $\mathbf{A}\hat{\mathbf{V}} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}$ always implies that $\mathbf{A}\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}\,\hat{\mathbf{V}}^H$. However, if $\hat{\mathbf{V}}$ is "only" a **nonsquare** matrix with orthonormal columns, the latter does not immediately allow us to conclude the reduced SVD, i.e. that $\mathbf{A} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}\,\hat{\mathbf{V}}^H$, since, it it is quite likely that $\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H \neq \mathbf{I}$. However, if we can append enough additional orthonormal columns to $\hat{\mathbf{V}}$ to create a square matrix, **and if** all of those additional columns lie in Null($\mathbf{A}$), then we can still derive the reduced SVD. To see why this is true, assume $\mathbf{V} = \left[\begin{array}{cc} \hat{\mathbf{V}} & \hat{\hat{\mathbf{V}}} \end{array}\right]$, where all the columns of $\hat{\hat{\mathbf{V}}}$ lie in Null($\mathbf{A}$). Show that, in this case

$$\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \mathbf{I} - \hat{\hat{\mathbf{V}}}\,\hat{\hat{\mathbf{V}}}^H$$

and therefore $\mathbf{A}\hat{\mathbf{V}}\,\hat{\mathbf{V}}^H = \mathbf{A}$.